Robustness of Interdependent Networks: The case of communication networks and the power grid

Marzieh Parandehgheibi and Eytan Modiano
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract—In this paper, we study the robustness of interdependent networks, in which the state of one network depends on the state of the other network and vice versa. In particular, we focus on the interdependency between the power grid and communication networks, where the grid depends on communications for its control, and the communication network depends on the grid for power. A real-world example is the Italian blackout of 2003, when a small failure in the power grid cascaded between the two networks and led to a massive blackout. In this paper, we study the minimum number of node failures needed to cause total blackout (i.e., all nodes in both networks to fail). In the case of unidirectional interdependency between the networks we show that the problem is NP-hard, and develop heuristics to find a near-optimal solution. On the other hand, we show that in the case of bidirectional interdependency this problem can be solved in polynomial time. We believe that this new interdependency model gives rise to important, yet unexplored, robust network design problems for interdependent networked infrastructures.

I. INTRODUCTION

Nowadays infrastructure networks are interdependent in such a way that the failure of an element in one network may cascade to another network and cause the failure of dependent elements. This failure procedure can cascade multiple times between two or more interdependent networks and result in catastrophic widespread failures. A particular example is the interdependency between the power grid and communication networks; where the 2003 Italian blackout, affecting the lives of fifty-five million people, was the result of such interdependency [1]. While reliability in networks has been studied extensively, the focus of most efforts has been on single networks in isolation. However, coupled networks exhibit very different behavior than isolated networks due to the cascading failure effects. In this paper, we develop a simple model for the interdependency between the networks, and analyze their robustness.

In order to enhance the reliability of the power grid, it is essential to have a smarter grid that detects autonomously faults and allows self-healing [2]. The control of today's power grid relies on a supervisory control and data acquisition (SCADA) system. This SCADA system does not have the capability to collect real-time data from the entire grid and analyze it efficiently to make control decisions. It is expected that future smart grid will need a wide-area measurement system (WAMS) which uses new technologies such as Phasor Measurement Units (PMUs) and Sensor and Actuator Networks (SANETs) to collect synchronous data, send it to control

centers for making online control decisions, and apply changes to the grid using devices such as actuators. This complex control network increases the dependency of the power grid on communication networks and gives rise to the necessity of analyzing and designing robust cyber-physical networks [3].

There have been a few attempts at introducing and modeling interdependent infrastructure networks. Rinaldi *et al.* describe interdependencies among major infrastructure networks including the power grid and communication networks. In particular, they show that all infrastructure networks depend on information delivered by the communication network for monitoring and controlling their subsystems (i.e. SCADA). At the same time, they show that communication networks depend on the physical output of other networks; for example, power from the electric grid for switches or water from the water infrastructure for cooling systems. These interdependencies are referred to as cyber and physical, respectively [4].

Later, Rosato *et al.* studied the Italian blackout of 2003 which affected the lives of fifty-five million people. They explain that this blackout was the result of a cascade of failures between the power grid and the communication network. In fact, due to failures in the power grid, some of the switches in the communication network lost their power and failed. Subsequently, due to failures in communication network, some of the substations in the power grid lost their control, and failed. They also demonstrate that failures inside the power grid lead to failures in the communication network using simulations on the real data of the Italian network [1]. Although the papers by Rinaldi and Rosato explain the concept of interdependency between networks, neither presents a mathematical model for investigating the behavior of interdependent networks. Moreover, there has been some work on the interdependency within layered communication infrastructure, e.g. between fiber optical networks and IP networks [5], [6]; but, they did not consider interdependency between different infrastructures.

In 2009, Buldyrev *et al.* presented a model for analyzing the robustness of interdependent random networks. They generate two disjoint random graphs and define a one-to-one dependency between every pair of nodes in these graphs. Using techniques from percolation theory, they show that random failures cascade through the graphs, and investigate the existence of a giant component [7]. Later, Parshani *et al.* showed that reducing the dependency between the networks makes them more robust to random failures [8]. A discussion

of follow-up works on Buldyrev's model can be found in [9].

However, real networks such as the power grid and the communication infrastructure are not random. These networks have known topologies; thus, it is essential to understand the impact of failure cascades in such networks in order to design robust interdependent infrastructures. To the best of our knowledge, this paper is the first attempt to model the interdependency between the networks with known topologies. The focus of this work is on the interdependency between power grid and communication networks; however, it can be extended to other networks with cyber-physical interdependency.

In Section II, we present the network model and define a new metric for assessing the robustness of networks. In Section III, we study the interdependency between networks with star topologies. In Section IV, we evaluate the robustness of a real network and discuss future research directions, and finally, conclude in Section V.

II. CYBER-PHYSICAL INTERDEPENDENCY MODEL

We start with a simple model for the power grid and the supporting Control and Communication Network (CCN). The power grid consists of generators G and substations S which are connected with power lines. Similarly, the CCN consists of control centers C and routers R which are connected with communication lines. Every router receives power from at least one substation and every substation sends data and receives control signals from at least one router. Figure 1 shows the interdependency model between the networks.
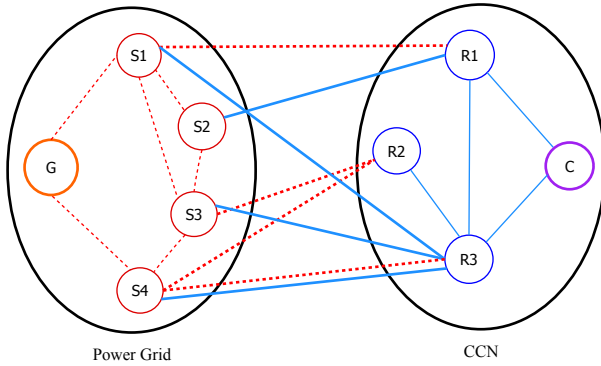

Fig. 1. Cyber-Physical Interdependency Model - dotted lines represent power lines and solid lines represent communication lines

In this model, we say that a substation *operates* if it has a path to a generator, i.e. receives power, and it is also connected to a router, i.e. sends data and receives control signals. Similarly, we say that a router *operates* if it has a path to a control center, i.e. sends data and receives control signals and it is also connected to a substation, i.e. receives power. Thus, a failure in the power grid may cause a failure in the CCN and vice versa. For example, future substations will be equipped with FACTS [1] which control the reactance of the power lines attached to a substation and prevent those lines from overloading and failure. These FACTS receive their

[1] Flexible Alternating Current Transmission System

control through the CCN; thus, the substations connected to a router will be controlled and do not experience power overloading. On the other hand, if a substation is not connected to a router (i.e. is not controlled), it may be overloaded and fail.

We assume that the power generators have internal control, and the control centers have backup generators; thus, they are robust to failures. In addition, in this paper, we do not consider the power flow equations, and assume that the connectivity of a substation to a generator is sufficient for receiving power. We also do not consider the amount of power supply or demand, and only focus on connectivity. Moreover, without loss of generality, we assume that there is only one generator and one control center; this is due to our model's assumption that every substation can be connected to any generator, and every router can be connected to any control centers. Although this model is not fully realistic, it captures the essential properties of interdependent networks.

A. Effect of a Single Failure

We start with an example demonstrating that a single failure can cascade multiple times within and between the power grid and CCN (Figure 2). Suppose that initially substation S_4 fails (Step 1). As a result, all the edges attached to S_4 fail, and router R_3 loses its power and fails (Step 2); Consequently, substations S_1 and S_3 lose their control, and router R_2 loses its connection to the control center C, and all fail (Step 3). Finally router R_1 loses its power, and substation S_2 loses its connection to the generator G, and both fail (Step 4).

In contrast, suppose that the power grid is not dependent on the CCN. In this case a substation fails, if and only if it is disconnected from the generator. For example, consider only the power grid from Figure 1 and, suppose that substation S_4 fails It can be seen that no other substation will be disconnected from the generator; thus, no further failure occurs. This example indicates that interdependency makes the networks more vulnerable, while it is essential to their operation.

It was seen in the example of Figure 2 that a single failure can lead to the failure of all nodes in both networks. We define the failure of all nodes in both networks as *Total Failure*. This observation leads to an important question. What is the minimum number of nodes whose removals will lead to total failure? We define the minimum total failure removals (MTFR) as a metric that helps to measure the robustness of a network, i.e. the larger the MTFR, the more robust the network. In particular, we consider two types of Node and Edge removals which we refer to as Node-MTFR and Edge-MTFR metrics.

In the case of a single network $G = \{V, E\}$, the smallest set of nodes that can cause a total failure, i.e. disconnect all the nodes in V from source S, is the set of nodes directly connected to S. Consequently, a star topology is the most robust network topology, since all of the nodes in a star are directly connected to the source.

(a) Step1 - S_4 fails, initially

(b) Step2 - R_3 fails

(c) Step3 - S_1, S_3 and R_2 fail
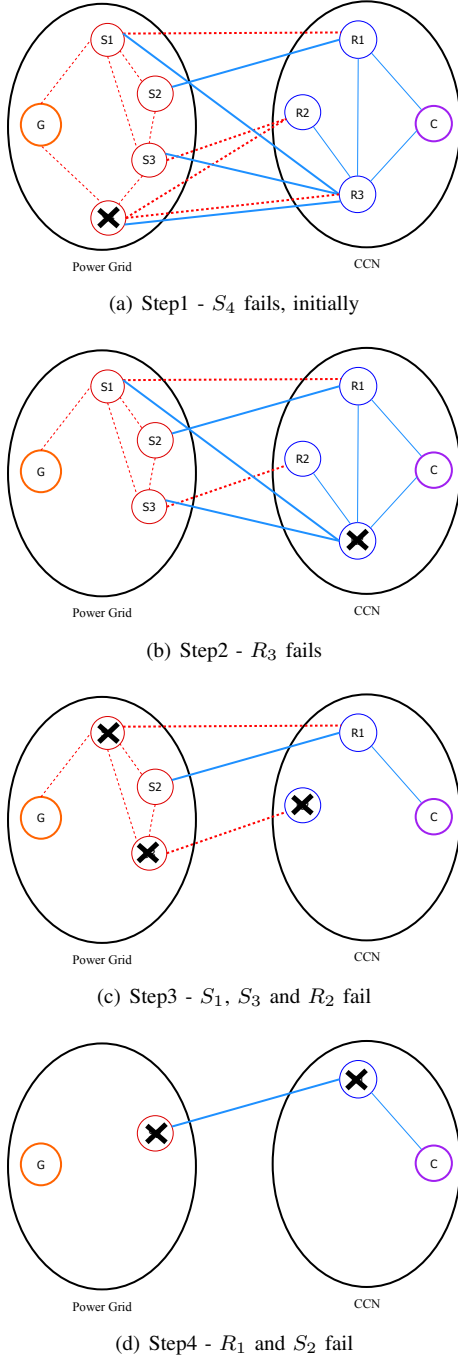

(d) Step4 - R_1 and S_2 fail

Fig. 2.   Cascade of a single failure in an interdependent model

III. INTERDEPENDENCY BETWEEN NETWORKS WITH STAR TOPOLOGIES

In order to analyze the interdependency between the power grid and the CCN, we assume that both networks have star topologies. Under this assumption, all of the substations in the power grid are directly connected to the generator; thus no substation's failure can disconnect the other substations from the generator. Similarly, all of the routers in the CCN are directly connected to the control center, and no router's

failure can disconnect the other routers from the control center. Therefore, any failure in the system would be *only* due to the interdependency between the networks, i.e. a substation fails if and only if it loses its connection to the CCN, and similarly a router fails if and only if it loses its connection to the power grid. This property gives us the opportunity to study the behavior of interdependent networks, and the effects of interdependency on the robustness of networks. Since both networks have star topologies, the interdependent network under study has a bipartite topology, i.e. no edge connects the nodes inside one network. In the following, we consider two distinct models of unidirectional and bidirectional interdependency. We analyze the networks under each model, and compare their complexity and robustness.

A. Unidirectional Interdependency

Under the unidirectional interdependency model, the edges between the networks are directed, i.e. if a power node S_i provides power for a router R_j, the router R_j does not necessarily provide the control signal for the same power node S_i (Figure 3(a)). In the following we show that finding the Node-MTFR is in fact a hitting cycle problem, and prove that it is an NP-complete problem. In order to do so we start with the following lemmas:
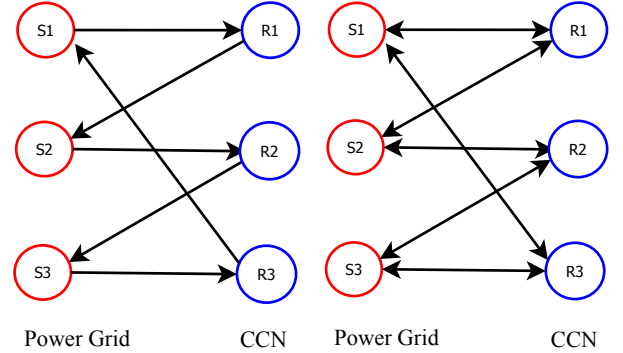

(a) Unidirectional Interdependency (b) Bidirectional Interdependency

Fig. 3.   Graph structure under different interdependency models

Lemma 1: A network with one or more operating nodes has at least one cycle.

Proof: We prove by contradiction that no node in a network can operate if there is no cycle in the network. Suppose that there is no cycle, i.e. all nodes are connected through one or more paths. First, remove all of the non-operating nodes; hence, the remaining nodes are operating, and the network is still acyclic. Now consider the starting node of one of the paths which is either a substation S_i or a router R_j. This starting node does not have any incoming edges; therefore, substation S_i (router R_j) does not receive any control (power) and cannot operate which is a contradiction with the assumption of nodes being operating.

Now we show that existence of at least one cycle is sufficient to have an operating node. In a bipartite graph, every

substation (router) in a cycle receives an incoming edge from a router (substation) in that cycle; thus, the nodes in that cycle can operate. If all the other nodes of the network receive an incoming edge directly or through a path starting from a node in that cycle, those nodes will be operating, too. ■

Lemma 2: To stop the operation of any cycle, one of the nodes in the cycle should be removed (Cycles are Stable Components).

Proof: By definition, every substation (router) in the bipartite graph remains operating if it has an incoming edge from a router (substation). Every node in a cycle receives at least one incoming edge from the nodes inside that cycle; therefore, the removal of nodes outside the cycle will not affect the operation of the nodes inside that cycle. As a result, to stop the operation of a cycle, one of its nodes must be removed. ■

Note that stopping the operation of a cycle is not equivalent to stopping the operation of all the nodes inside the cycle. If the nodes inside a cycle are isolated from the other nodes, removing exactly one node from the cycle will stop all of them from operating. However, if nodes inside a cycle receive incoming edges from other nodes outside of the cycle, more node removals are needed to cause the failure of all of the nodes in that cycle.

Lemma 3: For total failure, at least one node from every cycle should be removed .

Proof: By contradiction - Suppose that there exist a cycle so that none of its nodes are removed. By lemma 2, all of the nodes inside that cycle remain operating, which contradicts the assumption of total failure. ■

Theorem 1: The minimum number of nodes that hit all of the cycles in a bipartite graph is the optimal solution for the Node-MTFR problem.

Proof: Immediate from lemma 3. ■

Corollary 1: Finding the Node-MTFR in star networks with unidirectional interdependency is NP-complete.

Proof: By Theorem 1, the Node-MTFR problem is a hitting cycle problem which is exactly equivalent to the well-known problem of Feedback Vertex Set (FVS). By definition, FVS in a graph finds the smallest set of nodes so that their removals make the graph acyclic; and it is known to be NP-complete for general graphs [10]. Moreover, Cai *et al.* proved that FVS is NP-complete for a special class of bipartite graphs called bipartite tournament [11]. Therefore, the Node-MTFR problem which is finding FVS in general bipartite graphs is also NP-complete. ■

1) Problem Formulation: It was shown in lemma 3 that finding the Node-MTFR is a hitting cycle problem. Next, we present a cycle-based Integer Linear Programming (ILP) formulation for this problem, assuming that all of the cycles are given. Let N be a $n \times 1$ binary vector so that each component N_j takes values 1 if node j is removed and 0 otherwise. Let matrix $A \in R^{m \times n}$ be a mapping between the m cycles and n nodes, where $A_{ij} = 1$ if cycle i contains node j and $A_{ij} = 0$ otherwise. Let e be a $m \times 1$ vector of ones.

The Node-MTFR problem can be formulated as follows.

$$\text{minimize} \quad \sum_{j=1}^{n} N_j \quad (1)$$

$$\text{subject to} \quad A \times N \geq e \quad (2)$$

$$N_j \in \{0, 1\}, \quad j = 1, \cdots, n \quad (3)$$

In this formulation, the objective is to minimize the number of node removals. Every row i of constraint (2) requires that at least one of the removed nodes should hit cycle i. In the following, we develop heuristics to solve the problem.

2) Heuristics: Since computing the Node-MTFR is computationally difficult, we consider approximation algorithms that give a near-optimal set of node removals in polynomial time. As explained in section III-A, the node removals problem is equivalent to a hitting cycle problem. Thus, if we have the set of all cycles in the graph, we can apply a greedy algorithm devised for solving the hitting set problem. The input to the algorithm is the set of cycles (each cycle is defined as the set of nodes it contains), and the set of nodes in the graph. This cycle-based algorithm is an iterative algorithm that works as follows. In each iteration, it removes the node that is shared among maximum number of cycles, updates the set of cycles, and repeats until no cycle remains.

This cycle-based algorithm needs the set of all cycles as input; however, in general, a graph may have an exponential number of cycles. To overcome this deficiency, we devise a new algorithm that relies on the degree of the nodes instead of the cycles. The input to the algorithm is the adjacency matrix of the graph. The algorithm is iterative: Each iteration starts with a pruning stage in which the algorithm removes all of the edges that do not belong to a cycle. In the next stage of the iteration, it removes the node that has the maximum outgoing degree. Next, the algorithm removes all nodes that fail as a result of the cascading effect of that removal. Finally, the algorithm updates the adjacency matrix of the graph and repeats the iteration until no node remains.

In the following, we compare the performance of these algorithms with the optimal solution. We consider a random bipartite graph with N nodes on each side. Since enumerating all the cycles requires exponential time, we keep the size of N small, and limit the graph to have small cycles. To do that, instead of randomly generating edges, we randomly generate cycles of size 6 or smaller until all the nodes have at least one incoming edge. For each value of N, we generate 100 random graphs and then apply our algorithms to each graph in order to find the minimum node removals. Moreover, for the optimal solution, we solve the hitting cycle problem as given by (1)-(3) using CPLEX. As can be seen from Figure 4, on average, the degree-based algorithm gives a slightly larger number of nodes compared with the cycle-based algorithm and optimal solution; however, it is very fast as it does not need to enumerate all of the cycles.

3) Minimum Edge Removals: An alternative version of the problem is the minimum number of *edges* needed to be
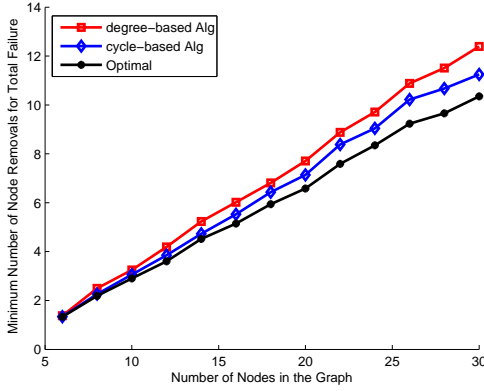
Fig. 4. Comparing different algorithms with optimal solution

removed to cause a total failure. Similar to lemmas 2 and 3, to stop the operation of any cycle, one should remove one of its edges, and to have a total failure, at least one edge from every cycle should be removed. Consequently, we have the following results:

Theorem 2: The minimum number of edges that hit all of the cycles is the optimal solution for the Edge-MTFR problem.

Proof: Trivial. ∎

Corollary 2: Finding the minimum edge removals for total failure in the star networks with unidirectional interdependency is NP-complete.

Proof: By Theorem 2, minimum edge removals problem is the edge version of the hitting cycle problem which is exactly equivalent to the well-known problem of Feedback Edge Set (FES). Similar to FVS, FES finds the smallest set of edges whose removals make the graph acyclic, and it is known to be NP-complete for general graphs [10]. Furthermore, Guo *et al.* proved that FES is NP-complete for bipartite tournaments [12]. Since finding FES in bipartite tournaments is a special case of the Edge-MTFR problem, the Edge-MTFR problem is also NP-complete. ∎

B. Bidirectional Interdependency

Under the bidirectional interdependency model, the edges between the two interdependent networks are bidirectional, i.e. if a power node S_i provides power to router R_j, router R_j *must* provide control to power node S_i (Figure 3(b)).

Theorem 3: Finding the minimum node removals for total failure in star networks with bidirectional interdependency is solvable in polynomial time.

Proof: It is easy to see from Figure 3(b) that edges are cycles of length two, and hitting cycles of length two guarantees hitting cycles of larger size. On the other hand, hitting at least one node in every cycle of size two is equivalent to finding the minimum vertex cover in bipartite graphs. By Konig's Theorem, finding the minimum vertex cover in bipartite graph is equivalent to maximum matching which is polynomially solvable [13]. Thus, finding the minimum node removals in star networks with bidirectional interdependency is polynomially solvable. ∎

For the Edge-MTFR, all of the edges must be removed. This is due to the fact that in the star networks with bidirectional interdependency every edge is a cycle, and cycles are robust components.

C. Comparing the interdependency models

We have seen that when networks have unidirectional interdependency, finding the optimal solution for the Node-MTFR problem is NP-complete; however, it can be solved in polynomial time when the networks have bidirectional interdependency. Here, we try to explain by way of an example why the analysis of unidirectional interdependency is more difficult than bidirectional interdependency. Figures 3(a) and 3(b) show two networks with the same topology under the different interdependency models. Suppose that in both networks, node S_1 is intentionally removed. It can be seen that the removal of S_1 in network 3(a) leads to the sequential failure of nodes R_1, S_2, R_2, S_3 and finally R_3. However, removal of S_1 in network 3(b) does not cause any failures, as all of the other nodes still have an incoming edge. Now, suppose that we want to cause the failure of node R_2 in network 3(b). In this case, all the nodes attached to R_2 (nodes S_1 and S_2) must be removed. It can be seen that failure of R_2 cannot cause the failure of other nodes as it is already disconnected from the rest of the network. These observations indicate that in the case of unidirectional interdependency, a failure can cascade multiple times between the networks. However, in the case of bidirectional interdependency, a failure cascades only in one stage: either from the power grid to the CCN or from the CCN to the power grid. This makes the analysis of bidirectional interdependent networks more tractable.

Next we compare the robustness of the interdependency models. We use the random graphs generated in section III-A2 to generate a new set of graphs with the same topology but bidirectional dependency. To compare the robustness of the two models, we find the optimal solution in the unidirectional graphs by solving the hitting set problem using CPLEX, and bidirectional graphs by solving the vertex cover problem. It can be seen from Figure 5 that for all values of N, networks with bidirectional dependency need more node removals; therefore, they are more robust to failures. This observation shows that the existence of more disjoint cycles and shorter cycles makes a network more robust.

IV. DISCUSSION

We measured the Node-MTFR metric on the Italian network topology shown in [1][2]. For the power grid, we only considered the substations that are directly connected to the generators. These substations are the most critical nodes as if they fail, all dependent substations will fail. Similarly, for the CCN, we only considered the routers that are directly connected to the control centers (we assumed that the control centers are located in the highly connected clusters of routers). Note that under this selection, the power grid and the CCN
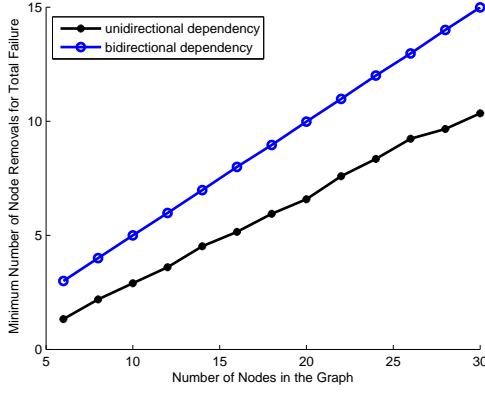
Fig. 5. Comparing the robustness of interdependency models

have star topologies. For the interdependency between the networks, we assumed that every substation receives its control signals from the nearest router, and every router receives power from the nearest substation. Figure 6 shows the power grid, the CCN and the interdependency between the two networks on the map of Italy.

Under the unidirectional interdependency model, we know that for a total failure, one should hit of all the cycles. It can be seen from Figure 6 that the cycles are very short, and mostly isolated. Therefore, one should remove a large number of nodes (13 nodes) simultaneously in order to cause the failure of the whole network. However, if we only consider the northwest of Italy, it can be seen that one third of the substations are controlled by a few routers. In fact, removing only three routers will lead to blackout in one third of the power grid. This shows that although the network might be robust to a total failure, only a few node removals can cause a considerable partial blackout. This observation leads to an important research direction: Measuring the effects of node removals on the failure of parts of the network.
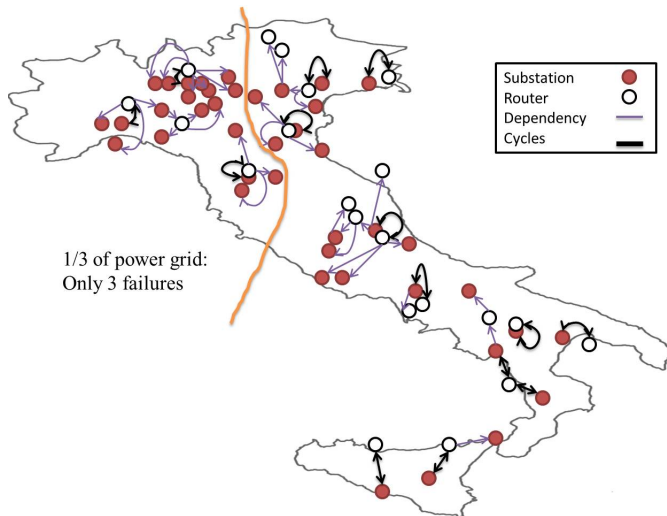

Fig. 6. Italian Power grid and CCN; Cycles shown in bold lines.

V. CONCLUSION

We considered the problem of robustness in interdependent networks. Two networks A and B are said to be interdependent if the state of network A depends on the state of network B and vice versa. We considered a cyber-physical interdependency between the networks in the sense that the nodes in network A depend on information from network B, and the nodes in network B depend on the physical output of network A. The massive blackout in Italy in 2003 was the result of such interdependency between the power grid and the CCN where a small failure in the power grid cascaded between the two networks and led to extensive failures in both the power grid and the CCN.

We studied the minimum number of nodes that should be removed from both networks so that all of the nodes in the networks fail after the ensuing cascades. We formulated this problem, and proved it is equivalent to the hitting cycles problem which is NP-complete in the case of unidirectional dependencies. We also presented polynomial algorithms which give suboptimal solutions. Moreover, for the case of bidirectional dependencies, we proved our problem is equivalent to the vertex cover problem in bipartite graphs which is solvable in polynomial time. We applied our results on the real network of Italy, and showed that it is robust to total failure. However, we showed that parts of the network are very vulnerable to failures.

REFERENCES

[1] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. Porcellinis, and R. Setola, "Modelling interdependent infrastructures using interacting dynamical models," *International Journal of Critical Infrastructures*, vol. 4, no. 1, pp. 63–79, 2008.
[2] M. Amin, "Toward self-healing energy infrastructure systems," *Computer Applications in Power, IEEE*, vol. 14, no. 1, pp. 20–28, 2001.
[3] E. Hossain, Z. Han, and V. Poor, Eds., *Smart Grid Communications and Networking*. Cambridge University Press, 2012.
[4] S. Rinaldi, J. Peerenboom, and T. Kelly, "Identifying, understanding, and analyzing critical infrastructure interdependencies," *Control Systems, IEEE*, vol. 21, no. 6, pp. 11–25, 2001.
[5] K. Lee and E. Modiano, "Cross-layer survivability in wdm-based networks," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 1017–1025.
[6] M. Parandehgheibi, H. Lee, and E. Modiano, "Survivable paths in multilayer networks," in *Information Sciences and Systems (CISS), 2012 46th Annual Conference on*. IEEE, 2012, pp. 1–6.
[7] S. Buldyrev, R. Parshani, G. Paul, H. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.
[8] R. Parshani, S. Buldyrev, and S. Havlin, "Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition," *Physical review letters*, vol. 105, no. 4, p. 48701, 2010.
[9] J. Gao, S. Buldyrev, H. Stanley, and S. Havlin, "Networks formed from interdependent networks," *Nature Physics*, vol. 8, no. 1, pp. 40–48, 2011.
[10] M. Garey and D. Johnson, *Computers and intractability*. Freeman San Francisco, CA, 1979, vol. 174.
[11] M. Cai, X. Deng, and W. Zang, "A min-max theorem on feedback vertex sets," *Mathematics of Operations Research*, vol. 27, no. 2, pp. 361–371, 2002.
[12] J. Guo, F. Hüffner, and H. Moser, "Feedback arc set in bipartite tournaments is np-complete," *Information processing letters*, vol. 102, no. 2, pp. 62–65, 2007.
[13] R. Ahuja, T. Magnanti, and J. Orlin, *Network flows: theory, algorithms, and applications*. Prentice hall, 1993.